

09/21/96

### Abstract of the Disclosure

A method and apparatus for extracting information from a file. A file is processed through an owning application with a printer driver to generate a representation having a modified format. Information is then extracted from the representation. The extracted information may include text strings and text characteristics of the text strings. The extracted information is then stored in a database. Local system processes or remote requestors may then search the database to locate files which include specific information. Using the text characteristic information, the located files are ranked according to the extent to which the specific information may be addressed within the located files.

5624

1992年12月